

CLAIMS

What is claimed is:

1) A computer program product comprising:

    a computer-usable medium having a computer program for transforming a source data structure to a destination data structure embodied therein, the computer program configured to:

        obtain a source data structure comprising a set of source tables, a set of source fields, a set of source records, a set of source table relationships, and a set of source values;

        obtain transformation data comprising information associated with at least one of the set of source tables, the set of source fields, the set of source records, the set of source table relationships, and the set of source values ;

        apply the transformation data to all or some of the set of source tables, the set of source fields, the set of source records, the set of source table relationships, or the set of source values;

        transform the source data structure into a destination data structure.

2) The computer program product of claim 1 wherein the destination data structure comprises a database.

3) The computer program product of claim 2 wherein the database comprises catalog data.

4) The computer program product of claim 2 wherein the database comprises financial data.

5) The computer program product of claim 1 wherein the transformation data comprises mapping information and the computer program is configured to use the mapping information to execute a means for mapping the source data structure to the destination data structure.

6) The computer program product of claim 5 wherein the mapping information is displayed to the user.

7) The computer program product of claim 5 wherein the mapping information comprises field-level mapping information that identifies a correlation between at least one of the set of source fields and at least one destination field.

8) The computer program product of claim 5 wherein the field-level mapping information is displayed to a user.

9) The computer program product of claim 7 wherein the at least one of the set of source fields in the field-level mapping comprises a source field combination having a plurality of source fields.

- 10) The computer program product of claim 9 wherein the source field combinations are displayed to a user.
- 11) The computer program product of claim 7 wherein the at least one destination field in the field-level mapping information comprises a destination field combination having a plurality of destination fields.
- 12) The computer program product of claim 11 wherein the destination field combinations are displayed to a user.
- 13) The computer program product of claim 7 wherein the at least one of the set of source fields in the field-level mapping information comprises a source field combination having a plurality of source fields and the at least one destination field in the field-level mapping comprises a destination field combination having a plurality of destination fields.
- 14) The computer program product of claim 13 wherein the source field combinations and the destination field combinations are displayed to a user.
- 15) The computer program product of claim 7 wherein the mapping information comprises value-level mapping information that identifies a correlation between at least one of the set of source values of the at least one of the set of source fields and at least one of the set of destination values of the at least one destination field.

- 16) The computer program product of claim 15 wherein the value-level mapping information is displayed to a user.
- 17) The computer program product of claim 15 wherein the at least one of the set of source values in the value-level mapping comprises a source value combination having the set of source values of a plurality of the at least one of the set of source fields associated with a set of destination values of the at least one of the destination fields.
- 18) The computer program product of claim 17 wherein the source value combinations are displayed to a user.
- 19) The computer program product of claim 15 wherein the value-level mapping comprises a destination value combination having a plurality of the set of destination values of the at least one of the destination fields associated with a plurality of the set of source values of the at least one of the set of source fields.
- 20) The computer program product of claim 19 wherein the destination value combinations are displayed to a user.
- 21) The computer program product of claim 15 wherein the set of source values of the at least one of the set of source fields in the value-level mapping comprises a source value combination having values of a plurality of the at least one of the set of source

fields and the destination values of the at least one of the destination fields in the value-level mapping comprises a destination value combination having a plurality of the at least one of the destination fields.

- 22) The computer program product of claim 21 wherein the source value combinations and the destination value combinations are displayed to a user.
- 23) The computer program product of claim 1 wherein the transformation data comprises type information and the computer program comprises a means for converting the set of source fields from a source type to a destination type based on the type information.
- 24) The computer program product of claim 23 wherein the type information is displayed to a user.
- 25) The computer program product of claim 1 wherein the transformation data comprises type information for converting the set of source fields from a source type to a destination type based on the type information.
- 26) The computer program product of claim 1 wherein the computer program is configured to merge the set of source values of the set of source fields into source value combinations comprising a plurality of source values and convert the source value combinations into destination fields of the destination data structure.

- 27) The computer program product of claim 26 wherein the source value combinations are displayed to a user.
- 28) The computer program product of claim 1 wherein the computer program is configured to generate hierarchy among the set of source values of the set of source fields into source value hierarchies comprising a plurality of source values and convert the source value hierarchies into destination fields of the destination data structure.
- 29) The computer program product of claim 28 wherein the source value hierarchies are displayed to a user.
- 30) The computer program product of claim 1 wherein the transformation data is generated automatically.
- 31) The computer program product of claim 1 wherein the transformation data comprises parsing information and the computer program is configured to execute a means for parsing data values from descriptive fields.
- 32) The computer program product of claim 1 wherein the source data structure comprises descriptive fields having data values and the computer program is configured to use the transformation data to extract the data values from the descriptive fields.

- 33) The computer program product of claim 1 wherein the computer program is configured to generate at least one added source field in accordance with the transformation data.
- 34) The computer program product of claim 33 wherein the added source field is displayed to a user.
- 35) The computer program product of claim 1 wherein the computer program is configured to generate at least one cloned source field containing a copy of the set of source values in one of the set of source fields in accordance with the transformation data.
- 36) The computer program product of claim 35 wherein the at least one cloned source field is displayed to a user.
- 37) The computer program product of claim 35 wherein the values of the at least one cloned source field is displayed to a user.
- 38) The computer program product of claim 1 wherein the computer program is configured to generate at least one split-into-hierarchy source field containing a hierarchy based on the set of source values in one of the set of source fields in accordance with the transformation data.

- 39) The computer program product of claim 38 wherein the at least one split-into-hierarchy source field is displayed to a user.
- 40) The computer program product of claim 38 wherein the set of source values of the split-into-hierarchy source field are displayed to a user.
- 41) The computer program product of claim 1 wherein the computer program is configured to generate at least one plurality of split-into-multiple source fields each containing components of the set of source values in one of the set of source fields in accordance with the transformation data.
- 42) The computer program product of claim 41 wherein the plurality of split-into-multiple source fields are displayed to a user.
- 43) The computer program product of claim 41 wherein the set of source values of the plurality of split-into-multiple source fields are displayed to a user.
- 44) The computer program product of claim 1 wherein the computer program is configured to extract data values from descriptive fields by identifying the data values within the descriptive fields, parsing the data values from the descriptive fields, and populating the at least one new source field with the data values.

45) The computer program product of claim 1 wherein the transformation data comprises measurement information and the computer program uses the measurement information to execute a means for normalizing units of measure within the set of source fields.

46) The computer program product of claim 45 wherein the measurement information is displayed to a user

47) The computer program product of claim 45 wherein the means for normalizing further comprises combining numeric value and the unit of measure from a plurality of source fields.

48) The computer program product of claim 1 wherein at least one of the set of source values within the set of source fields comprises an improperly formed measurement value comprising a numeric value and a unit of measure.

49) The computer program product of claim 48 wherein the unit of measure comprises an implicit unit of measure associated with the set of source values .

50) The computer program product of claim 48 wherein the unit of measure comprises inconsistent string values.

- 51) The computer program product of claim 48 wherein the units of measure differ within a plurality of the values within each of the set of source fields.
- 52) The computer program product of claim 48 wherein computer program is further configured to append the unit of measure to set of source values missing the unit of measure.
- 53) The computer program product of claim 1 wherein the source data structure comprises a plurality of joined data sources.
- 54) The computer program product of claim 1 wherein the transformation data comprises matching information and the computer program is configured to use the matching information to execute a means for matching the set of source records to the destination records.
- 55) The computer program of claim 54 wherein the matching information comprises record-level information identifying the correlation between at least one of the set of source records and at least one of the destination records.
- 56) The computer program product of claim 54 wherein the matching information is displayed to a user.

- 57) The computer program product of claim 54 wherein the matching information indicates a new destination record is to be created with at least one of the set of source values from one of the set of source fields from one of the set of source records.
- 58) The computer program product of claim 54 wherein the matching information indicates at least one of the destination fields in at least one of the destination records is to be updated with at least one of the set of source values from one of the set of source fields.
- 59) The computer program product of claim 54 wherein the matching information indicates at least one destination record is to be replaced with at least one of the set of source records.
- 60) The computer program product of claim 1 wherein the computer program is further configured to transform the set of source values in accordance with the transformation data.
- 61) The computer program product of claim 60 wherein the transformed set of source values are displayed to a user.
- 62) The computer program product of claim 1 wherein the computer program comprises an integrated interface for obtaining the transformation data.

- 63) The computer program product of claim 62 wherein the source data structure is represented within the integrated interface as a hierarchy.
- 64) The computer program product of claim 63 wherein the hierarchy comprises a visual representation of the set of source tables.
- 65) The computer program product of claim 63 wherein the hierarchy comprises a visual representation of the set of source fields.
- 66) The computer program product of claim 63 wherein the hierarchy comprises a visual representation of the source table relationships.
- 67) The computer program product of claim 66 wherein a user can define additional relationships between the set of source tables.
- 68) The computer program product of claim 63 wherein the hierarchy comprises a visual representation of the source data values.
- 69) The computer program product of claim 62 wherein the destination data structure is represented within the integrated interface as a hierarchy.
- 70) The computer program product of claim 69 wherein the hierarchy comprises a visual representation of the destination tables.

- 71) The computer program product of claim 69 wherein the hierarchy comprises a visual representation of the destination fields.
- 72) The computer program product of claim 69 wherein the hierarchy comprises a visual representation of the destination table relationships.
- 73) The computer program product of claim 69 wherein the hierarchy comprises a visual representation of the destination data values.
- 74) The computer program product of claim 1 wherein the computer program comprises a means for handling exceptions a source field at a time.
- 75) The computer program product of claim 1 wherein the computer program is configured to collapse the set of source values within the set of source records down to a set of distinct values within the set of source records.
- 76) The computer program product of claim 75 wherein the set of distinct values is configured to act as a proxy for the set of source values.
- 77) The computer program product of claim 75 wherein each distinct value within the set of distinct values is configured to act as a proxy for all instances of the distinct value across the set of source records.

78) The computer program product of claim 75 wherein the transformation data is applied to the set of distinct values.

79) The computer program product of claim 75 wherein the transformation data is applied once to each distinct value rather than once for each instance of the distinct value, and is automatically propagated to each instance of the distinct value.

80) A computer program product comprising:  
a computer-useable medium having a computer program for transforming a source structure to a destination data structure embodied therein, the computer program configured to:  
obtain a source data structure;  
obtain transformation data comprising information associated with the source data structure;  
collapse the set of source values within the set of source records down to a set of distinct values within the set of source records;  
apply the transformation data to the set of distinct values;  
transform the source data structure into the destination data structure.

81) The computer program product of claim 80 wherein the source data structure comprises a structured document.

82) The computer program product of claim 80 wherein the destination data structure comprises a structured document.

83) An apparatus for transforming data comprising:

- a processor;
- a memory medium coupled to the processor;
- the memory medium containing a computer program configured to present a graphical user interface comprising:

  - a first region comprising a source hierarchy representing a source data structure;
  - a second region comprising a destination hierarchy representing a destination data structure;
  - a third region configured to obtain transformation data associated with the source data structure and destination data structure.

84) The apparatus of claim 83 wherein the transformation data comprises partitioning information.

85) The apparatus of claim 83 wherein the transformation data comprises field-level mapping information.

86) The apparatus of claim 83 wherein the transformation data comprises value-level mapping information.

- 87) The apparatus of claim 83 wherein the transformation data comprises matching information.
- 88) The apparatus of claim 83 wherein the transformation data comprises type conversion information.
- 89) The apparatus of claim 83 wherein the transformation data comprises parsing information.
- 90) A computer program product comprising:
  - a computer-readable medium having a computer program for transforming a source structure to a destination data structure embodied therein, the computer program configured to:
    - obtain the source data structure;
    - obtain transformation data comprising information associated with the source data structure;
    - collapse the set of source values within the set of source records down to a set of distinct values within the set of source records wherein each distinct value within the set of distinct values is configured to act as a proxy for instances of the distinct value across the set of source records;
    - apply the transformation data to the set of distinct values;
    - transform the source data structure into the destination data structure.